Genomic occupancy of the bromodomain protein Bdf3 is dynamic during differentiation of African trypanosomes from bloodstream to procyclic forms

Ethan Ashby[2], Lucinda Paddock[1], Hannah L. Betts[1], Jingwen Liao[1], Geneva Miller[1], Anya Porter[1], Lindsey M. Rollosson[1], Carrie Saada[1], Eric Tang[1], Serenity J. Wade[1], Johanna Hardin[2], Danae Schulz[1*]

1.) Department of Biology, Harvey Mudd College, 301 Platt Boulevard, Claremont, CA 91711

2.) Department of Mathematics and Statistics, Pomona College, 333 N College Way, Claremont, CA 91711

*Corresponding author, dschulz@g.hmc.edu

Abstract

Trypanosoma brucei, the causative agent of Human and Animal African trypanosomiasis, cycles between a mammalian host and a tsetse fly vector. The parasite undergoes huge changes in morphology and metabolism during adaptation to each host environment. These changes are reflected in the differing transcriptomes of parasites living in each host. However, it remains unclear whether chromatin interacting proteins help mediate these changes. Bromodomain proteins localize to transcription start sites in bloodstream parasites, but whether the localization of bromodomain proteins changes as parasites differentiate from bloodstream to insect stages remains unknown. To address this question, we performed Cleavage Under Target and Release Using Nuclease (CUT&RUN) against Bromodomain Protein 3 (Bdf3) in parasites differentiating from bloodstream to insect forms. We found that Bdf3 occupancy at most loci increased at 3 hours following onset of differentiation and decreased thereafter. A number of sites with increased bromodomain protein occupancy lie proximal to genes with altered transcript levels during differentiation, such as procyclins, procyclin associated genes, and invariant surface glycoproteins. Most Bdf3 occupied sites are observed throughout differentiation. However, one site appears *de novo* during differentiation and lies proximal to the procyclin gene locus housing genes essential for remodeling surface proteins following transition to the insect stage. These studies indicate that occupancy of chromatin interacting proteins is dynamic during life cycle stage transitions, and provides groundwork for future studies on the effects of changes in bromodomain protein occupancy. Additionally, the adaptation of CUT&RUN for *Trypanosoma brucei* provides other researchers an alternative to chromatin immunoprecipitation (ChIP).

Importance

The parasite *Trypanosoma brucei* is the causative agent of Human and Animal African Trypanosomiasis (sleeping sickness). Trypanosomiasis, which affects humans and cattle, is fatal if untreated. Existing drugs have significant side effects. Thus, these parasites impose a significant human and economic burden in Sub-Saharan Africa where trypanosomiasis is endemic. *T. brucei* cycles between the mammalian host and a tsetse fly vector, and parasites undergo huge changes in morphology and metabolism to adapt to different hosts. Here, we show that DNA-interacting Bromodomain Protein 3 (Bdf3) shows changes in occupancy at its binding sites as parasites transition from the bloodstream to the insect stage. Additionally, a new binding site appears near the locus responsible for remodeling of parasite surface proteins during transition to the insect stage. Understanding the mechanisms behind host adaptation is important for understanding the life cycle for the parasite.

Introduction

The ability to adapt to different environments is vital for parasites that live in different hosts. *Trypanosoma brucei*, the protozoan parasite that causes Human and Animal African Trypanosomiasis, is one such organism. *T. brucei* lives in the bloodstream and tissues of the mammalian host (1, 2) and multiple organs within the tsetse fly vector as it travels from the gut to the salivary gland (3). Throughout its life cycle, the parasite adapts to each unique environment. Parasites living in the bloodstream of the mammal evade the host immune system through antigenic variation of surface proteins called Variant Surface Glycoproteins (VSGs) (4, 5). Prior to transitioning to the fly, bloodstream parasites differentiate to stumpy forms that are transcriptionally pre-adapted for making the transition to the fly gut (6). Once the parasites arrive in the midgut, they differentiate fully to procyclic forms, and VSGs on the surface are replaced by procyclin proteins (7). Because the mammalian bloodstream and fly midgut differ in temperature, pH, and nutrient availability, the parasites undergo huge changes in morphology and metabolism. Underlying these changes are large differences in transcript levels for thousands of genes (8–11). Researchers have made great strides in understanding how environmental signals are sensed by the parasites and what signaling pathways may be important for the transition to stumpy and procyclic forms (12–18). It has also been demonstrated that RNA binding proteins play a role in differentiation processes (19–25). However, whether chromatin interacting proteins play a role in initiating or regulating changes in transcript levels necessary for transition from the bloodstream to the procyclic stage in *T. brucei* is less well understood.

Experimental observations suggest that chromatin interacting proteins might be involved in transcriptome reprogramming during the transition from bloodstream to procyclic forms. Notably, inhibition of chromatin interacting bromodomain proteins in bloodstream parasites results in changes to the transcriptome that mirror those that occur as parasites transition from the bloodstream form to the procyclic form (26). Bromodomain proteins bind to acetylated histone tails in *T. brucei* and other model organisms (27–30) and have well established roles in gene regulation. Proteins with bromodomains have been shown to play a number of gene regulatory roles, including histone modification, chromatin remodeling, transcription factor recruitment, and enhancer or mediator complex assembly (31). They have been shown to be activators of gene transcription in some contexts, but can also be involved in gene silencing (32). Degradation of the mammalian Brd4 bromodomain protein in a leukemia cell line model results in a global disruption of productive transcription elongation driven by collapse of the elongation complex (33). Inhibition of bromodomain proteins in mammalian stem cells results in spontaneous differentiation (34, 35). Finally, mammalian Brd4 has been shown to be recruited to lineage specific enhancers (36) and is necessary for adipogenesis and myogenesis (37). In *T. brucei*, seven bromodomain proteins (Bdfs) have been identified. Six of these proteins bind to transcription start sites at areas where polycistronic transcription units diverge (26, 29, 38). Members of the bromodomain protein family form distinct complexes in bloodstream forms (38). The focus of this study is Bdf3, which localizes to transcription start sites and has been shown to associate with Bdf5 and the Histone acetyltransferase HAT2 (38). Bdf3 was chosen as a good first candidate because knockdown of this protein by RNAi results in transcriptome changes similar to those that occur during differentiation from bloodstream to procyclic forms (26). While the localization of Bdf3 has been well characterized in bloodstream parasites, nothing is known about whether this localization is maintained during differentiation to procyclic forms. We hypothesized that Bdf3 might undergo changes in localization or occupancy at binding sites during differentiation from bloodstream to procyclic forms. Such changes in occupancy or localization could play a role in transcriptome reprogramming during differentiation.

To investigate whether localization of Bdf3 is dynamic during the transition from bloodstream to procyclic forms, we analyzed differentiating parasites using Cleavage Under Targets and Release Using Nuclease (CUT&RUN). This technique is an alternative to Chromatin Immunoprecipitation and sequencing (ChIP-seq) that avoids potential artifacts and can be performed in less time (39, 40). Our results indicate that the CUT&RUN protocol developed in mammalian systems can be modified for successful use in *T. brucei*. We were able to use data from differentiating parasites processed by CUT&RUN to show that 3 *de novo* sites of Bdf3 localization appear near the procyclin gene locus in differentiating parasites. More globally, occupancy at the majority of Bdf3 binding sites is transiently increased during the course of differentiation from bloodstream to procyclic forms.

Results and Discussion

Optimization of CUT&RUN for bloodstream form *T. brucei*

The CUT&RUN protocol was originally developed for use in mammalian cell systems (39), so we set out to adapt the protocol to *T. brucei* bloodstream parasites. In brief, the CUT&RUN protocol works as follows. Cells are permeabilized and incubated with an antibody against the protein of interest. A fusion protein of protein A and micrococcal nuclease (pA-MN) is then added. The fusion protein binds to the antibody, and when calcium is added, the fusion protein cleaves the DNA immediately surrounding the protein of interest. The small DNA fragments that are released diffuse out of the nucleus and can be collected in the supernatant following centrifugation of the permeabilized cells. These small DNA fragments are used to generate a sequencing library, which can be analyzed with statistical methods already developed for ChIP-seq. When CUT&RUN is performed using an antibody against an abundant histone protein, a characteristic ladder of bands is produced that represent 150bp increments corresponding to the size of DNA wrapped around a nucleosome (39, 41). Thus, we optimized the CUT&RUN protocol in the *T. brucei* EATRO1125 Antat 1.1 pleomorphic parasite line using an antibody against the abundant protein histone H3.

To optimize the permeabilization step of CUT&RUN, we developed a flow cytometry assay to test permeabilization of parasite membranes. Parasites were permeabilized with various detergents and then incubated with a primary antibody against histone H3. A fluorescently labeled secondary antibody was then added and the parasites were assayed by flow cytometry. We observed anti-H3 staining following permeabilization with saponin, but did not see this staining in control samples where anti-H3 was not added (Fig.1A). Having optimized the permeabilization conditions, we next tested a series of incubation times and temperatures for the protein A micrococcal nuclease cleavage step of the protocol. We found that incubation for 5 minutes at 37ºC most efficiently produced the characteristic ladder of bands expected following successful cleavage around histone H3 (Fig. 1B). However, because of the danger of non-specific cleavage at 37ºC (39), we processed experimental CUT&RUN samples for 5 minutes at 25ºC. Targeted micrococcal nuclease cleavage was dependent on the addition of calcium and did not occur when a non-specific control antibody was used (Fig. 1C). These results indicate that we successfully developed a CUT&RUN protocol for use in *T. brucei* bloodstream parasites.

Bdf3-HA tagged *T. brucei* parasites express a stumpy induction marker when grown to high density

To study the function of Bdf3 in differentiation competent parasites, we generated EATRO1125 Antat 1.1 pleomorphic parasites with an HA-tagged allele of *BDF3* and knocked out the

remaining *BDF3* allele (Fig. 2A). Correct targeting to the endogenous *BDF3* locus was verified using a PCR assay (Fig. 2A, 2B) and our experiments were conducted with clone 6 of 24 tested clones (for simplicity, some clones are omitted from the figure). We tested our tagged parasites for their ability to generate stumpy forms by growing them to high density and measuring transcript levels of *PAD1*. We found increased transcript levels of *PAD1* when our Bdf3-HA tagged parasites were grown to high density as compared to parasites grown at low density (Fig. 2C). This accords with previous work showing increased *PAD1* transcript levels in pleomorphic parasites grown to high density (13), and indicates that our Bdf3-tagged pleomorphic strain shows the expected increase in the Pad1 stumpy induction marker once stumpy formation is induced.

Bdf3 binding sites identified by CUT&RUN are similar to those identified by ChIP-seq in bloodstream parasites

Once we optimized the protocol for CUT&RUN in bloodstream parasites, we used the pleomorphic HA-tagged Bdf3 strain described above to perform CUT&RUN in bloodstream parasites using an anti-HA antibody. A non-specific IgG antibody was used as a control. We used MACS software (42) to call peaks of Bdf3 localization using the IgG sample as a control. We compared peaks found by CUT&RUN to our published Bdf3 peaks found during ChIP-seq (26) and found that peaks called by MACS for our CUT&RUN dataset were very similar to peaks identified in our ChIP-seq dataset (Fig. 3A). 92% of peaks identified by ChIP-seq were also identified by CUT&RUN (Fig. 3B) and 86% of CUT&RUN peaks were identified by ChIP-seq. Overall this suggests that CUT&RUN is a viable technique for producing localization data in *T. brucei*.

Because CUT&RUN does not use formaldehyde crosslinking, this technique could avoid artifacts that have been previously documented for ChIP-seq (40). In CUT&RUN, binding of the primary antibody to the target protein of interest occurs following permeabilization and prior to any other processing. The data produced for Bdf3 localization by CUT&RUN is quite similar to what has been seen previously for the localization of Bdf3 in bloodstream forms using ChIP-seq (Fig. 3B). Bdf3 was found to localize primarily to divergent strand switch regions thought to be transcription start sites. 76% of all divergent strand switch regions (Table S7) (112 out of 148) were within 5kb of a Bdf3 peak. This is in accord with previous results that show Bdf3 localizing to regions where transcription is initiated (26, 29, 38). In addition, peaks of localization called by MACS showed substantial overlap between the two techniques. This is reassuring, as it suggests that data acquired by ChIP-seq represents physiological levels of binding for the *T. brucei* proteins that have been studied using the ChIP-seq technique. One situation where CUT&RUN might fail is in detecting transient protein-DNA interactions. For protein interactions of this type, a crosslinking-based technique might be necessary and/or preferable.

Although the data produced by ChIP-seq and CUT&RUN are similar, CUT&RUN is faster. ChIP-seq requires at least two overnight steps to immunoprecipitate complexes and reverse the formaldehyde crosslinking, with a potential third overnight step to bind primary antibodies to beads. These steps are then followed by sequencing library processing. In contrast, CUT&RUN can be performed in ~4 hours. We were able to perform CUT&RUN experiments successfully using 50 million cells, while 100 million cells is generally recommended for ChIP-seq protocols in *T. brucei.* It is possible that even fewer cells could be used successfully, but we did not formally test this. The originators of the protocol in mammalian cells claim that high quality data can be obtained using 100-1000 cells (43). The development of CUT&TAG (Cleavage Under Targets and TAGmentation) has streamlined the process even further by eliminating some downstream steps associated with sequencing library preparation (43), and has further been

refined to be fully automated (44). CUT&TAG has also been adapted for single cell chromatin studies (45), which represents an exciting future application for the technique in *T. brucei*. A recent paper used single cell sequencing to delineate transcriptome changes that occur during differentiation (8). The use of CUT&TAG on single cells during differentiation might delineate the accompanying changes in occupancy for chromatin associated proteins throughout the differentiation process. Excitingly, CUT&RUN has been successfully adapted for use in *Toxoplasma gondii* to identify a master regulator of differentiation (46).

Bdf3 localizes to a region near the procyclin gene locus following differentiation.

In order to ascertain whether the localization of Bdf3 is altered as parasites differentiate from the bloodstream to the procyclic form, we induced differentiation of our pleomorphic Bdf3-HA tagged strain by resuspending parasites in differentiation media, adding 6mM cis-aconitate, and incubating them at 27ºC. We harvested parasites in triplicate at 1h, 3h, 24h, and 76h post differentiation and performed CUT&RUN using an anti-HA antibody. Bloodstream parasites were also processed the same way using 5 biological replicates. Sequencing libraries were generated and the resulting reads were trimmed for quality and aligned to the *T. brucei* genome. MACS (42) was then used to call peaks of Bdf3 localization at every time point (Fig. 4, peak locations given in Table S1, Table S8, Table S9). Visual inspection of the results revealed that almost all Bdf3 peaks identified in all 5 replicates of bloodstream parasites were also identified as peaks at every timepoint thereafter (Fig. 4).

While most MACS identified Bdf3 peaks were retained throughout the course of differentiation, one notable exception to this trend was found at chromosome 10 in the region of the procyclin gene locus (47). MACS called three Bdf3 peaks in this region at 76 hours post differentiation that were not identified as peaks earlier in the time course (Fig. 5). Quantification of Bdf3 occupancy at this site using the DiffBind program (48) revealed that Bdf3 occupancy increases at 3h and remains high throughout the remainder of the timecourse (Fig. S1A) We tested *EP1* transcript levels following differentiation in our pleomorphic line using the same timepoints and found a ~ 5 fold increase in *EP1* transcript levels at 1h and a ~20-50 fold increase at 3 hours (Fig. S2). While the MACS algorithm identified Bdf3 peaks at 3 specific sites (indicated by the peaks labeled as 520,521, and 523), the gene track for Bdf3 shows widespread binding over a large region at this locus (Fig. 5B). This is especially interesting because transcripts from the *EP* and *PAG* genes near this locus are increased during the transition from bloodstream to procyclic forms so that parasites can remodel their VSG surface coat with procyclin. In contrast to many genes driven by Pol II, procyclin genes are instead transcribed by Pol I (49–52). Additionally, while pol II driven genes are thought to be largely regulated post-transcriptionally, procyclin genes have been shown to be transcriptionally regulated (53). While RNA binding proteins have been shown to have a role in stabilizing transcripts of procyclin genes (23–25), the *de novo* localization of Bdf3 to this region may indicate that bromodomain proteins play a role in inducing or maintaining increased levels of procyclin transcripts during differentiation (Fig. 5). This idea is in accord with findings that chromosomal context is important for procyclin promoter regulation (54). This would be interesting to test in future studies, perhaps through the use of a tethering experiment where Bdf3 is artificially localized to the *EP* locus in bloodstream forms. An increase in the level of transcript for *EP1* following artificial tethering would support a model where the *de novo* appearance of Bdf3 at this locus helps to increase transcript levels of procyclin genes nearby, facilitating surface remodeling of the parasite.

Custom scripts were used to systematically check for other instances of *de novo* Bdf3 peak formation following differentiation. While a number of sites were identified, visual inspection revealed that most of these sites were likely false positives. Many were in regions with high background in the IgG control samples. Thus, it's likely that for some timepoints, these regions

just made it over the threshold to be called by MACS as a peak, while at other timepoints they did not, leading to a false positive result. Table S2 lists each of the sites with a numerical code given by visual inspection, with 0 indicating no evidence of *de novo* peak formation by visual inspection, 1 indicating some evidence, and 2 indicating good evidence. Sites that received a score of 1 include the *GPEET* (Tb927.6.510) procyclin locus on chromosome 6 (Fig. S6) and the Invariant Surface Glycoprotein (*ISG75)* locus on chromosome 5 proximal to Tb927.5.360 (Fig. S3). Both of the protein products for these loci are associated with differentiation processes. A similar analysis was performed to check for MACS-called Bdf3 peaks present in bloodstream forms that were not present at subsequent time points. Again, many of these loci likely represent false positives, and only two sites were scored as 1: a *VSG* locus on chromosome 5 and an *ESAG* locus on chromosome 1 (Table S3).

Occupancy of Bdf3 at genomic binding sites transiently increases as parasites differentiate from bloodstream to procyclic forms

Having ascertained that most sites of Bdf3 localization are retained throughout differentiation, we next wanted to measure whether occupancy at these sites is altered by quantifying tag counts within each Bdf3 peak over time. To do so, we took advantage of the DiffBind program (48, 55) which is designed to identify changes in protein occupancy under different conditions. DiffBind takes BAM files and MACS called peaks as input, and first finds consensus peaks that are present in all biological replicates using the provided MACS files. These consensus peaks are considered regions of interest. The program normalizes the data and computes a normalized tag count for regions of interest using a 400bp region surrounding the peak summit. A region is considered to have a change in occupancy if there is a statistically significant change in tag count for the region around the peak summit at a time point after differentiation as compared to the bloodstream samples.

The method of normalization has been shown to have an outsized effect on whether a particular region is identified as having a significant change in occupancy (48). To circumvent this potential issue, we used four different normalization methods to identify regions with a change in occupancy over the course of differentiation. The regions identified as having a change in occupancy by all four normalization methods are considered ‘high confidence’ differentially occupied regions. The normalization methods used include (1) reads per kilobase of transcript per million mapped reads (RPKM), (2) spike-in library size normalization, which normalizes using spiked in yeast DNA for each sample (3) background RLE (Relative Log Expression), where counts are divided by sample-specific size factors determined by median ratio of gene counts relative to geometric mean per gene, a method similar to DESeq, and (4) background RLE of spiked in reads, which uses the former method with spiked-in yeast in reads. In total, we found 268 ‘high confidence’ regions that were identified as having a significant change in occupancy at a time point following differentiation when compared to the bloodstream data (Fig. 6).

We plotted the normalized tag counts for each of the 268 ‘high confidence’ differentially occupied Bdf3 binding sites arrayed by chromosome (Fig. 7A, Fig. S4, Table S4). There was a remarkable similarity to the pattern of occupancy for these differentially bound regions over time, with most regions showing a peak in occupancy at 3 hours post differentiation and a decrease in occupancy thereafter. In order to ensure that the peak in occupancy observed for Bdf3 at 3 hours post differentiation was not an artifact, we randomly shuffled Bdf3 binding site locations using bedtools (56). The regions were shuffled such that each shuffled site was retained on the same chromosome as the original site and all regions identified as consensus sites were excluded. These shuffled regions were then run through the DiffBind program and

normalized the same way as the Bdf3 consensus sites (Fig. 7B, Fig. S5, Table S5). When tag counts for shuffled control regions were plotted over time, we did not observe the same peak in the number of tag counts at the 3-hour time point, indicating that the peak in Bdf3 occupancy that occurs 3 hours after differentiation represents a genuine change in the amount of Bdf3 found at that location, either at the single cell or the population level. Interestingly, the occupancy for peaks at the *EP* locus does not show this pattern. Instead, occupancy shows a pronounced increase at 3h and remains high for the duration of the timecourse (Fig. S1). The fact that there is a pronounced increase in *EP1* expression at 3h (Fig. S2) may indicate that the *de novo* formation of Bdf3 peaks at this time point could help increase transcript levels of *EP1* at this locus. While the MACS algorithm did not call Bdf3 peaks at this locus until 72 hours, quantification by Diffbind as well as visual inspection of the sequencing tracks indicate that the peaks may be forming as early as 1-3 hours (Fig. S1, Fig 5).

The fact that the peak of Bdf3 occupancy is reached 3h following addition of cis-aconitate is especially interesting in light of work showing that commitment to differentiation occurs 2-3h hours after cis-aconitate treatment (57). This group further demonstrated that protein synthesis is required to generate 'memory' of exposure to the differentiation signal. One possible model is that bromodomain protein occupancy at transcription start sites facilitates commitment at 3h following differentiation by an unknown mechanism. For instance, if increased occupancy led to an increase in transcript levels, this could aid in the protein synthesis required for commitment to differentiation.

Of the 268 'high confidence' sites with changes in Bdf3 occupancy, 162 are within 5kb of a divergent strand switch region considered to be sites of transcription initiation in *T. brucei*. This might be an underestimate of the true number of dynamically occupied Bdf3 sites at divergent strand switch regions because we used only unique alignments, and thus we lack data for highly repetitive regions with divergent strand switch regions near the ends of the chromosomes. Of 148 divergent strand switch regions we examined, 60% were within 5kb of a 'high confidence' dynamically occupied Bdf3 site. The effect of the increased occupancy at sites of Bdf3 localization is not yet known for *T. brucei*, but in other systems, changes in occupancy for chromatin binding proteins can reflect an underlying change in particular histone tail modifications that are bound by the chromatin binding protein (55). In our particular case, a change in histone acetylation at transcription start sites could result in the observed increase in Bdf3 occupancy at these sites as parasites begin differentiation. Future studies could investigate the acetylation levels at transcription start sites after the induction of differentiation from the bloodstream to the procyclic form. A transient increase in acetylation at transcription start sites might result in a corresponding transient increase in bromodomain protein localization at these sites that is observed for Bdf3 (Fig. 7). This leaves us with the question of how an increase in acetylation might support the differentiation process. One model is that such an increase in acetylation might increase genomic transcription overall, and thus help the parasites exit the relatively quiescent cell cycle arrested stumpy stage and transition to the dividing procyclic stage. In other systems, quiescence results in a global decrease in acetylation levels (58–60). Specific chromatin remodeling enzymes are required to generate hypertranscription necessary for exiting the quiescent state (61). If an increase in acetylation levels and increased transcription are necessary during the transition from the quiescent stumpy form to the cycling procyclic form, this could result in the observed increase in Bdf3 occupancy following the differentiation cue. Bromodomain protein mediated regulation of global Pol II transcript levels has been demonstrated in *Leishmania* (62). RNA-seq analyses are typically normalized in such a way as to obscure global increases or decreases in transcript levels (63). Thus, it might be interesting to examine whether the exit from the stumpy to the procyclic stage results in a global increase in transcript levels for Pol II regulated genes.

Bromodomain inhibition in bloodstream parasites using the small molecule inhibitor I-BET151 results in changes in the transcriptome that mirror those that occur during differentiation from the bloodstream to the procyclic stage in many ways, including an increase in *EP1* transcript levels (26). We observed a sharp decrease in bromodomain protein occupancy for at least 268 Bdf3 binding sites in the period between 3h and 24h after the initiation of differentiation (Fig. 7). One model for why bromodomain inhibition may trigger transcriptome changes akin to differentiation is that binding of the drug to its bromodomain protein target causes a sharp decrease in occupancy for bromodomain proteins in bloodstream forms as is seen at 3 hours post differentiation (Fig. 7). We did observe decreased enrichment of Bdf3 in I-BET151 treated bloodstream parasites at several sites, consistent with this model (26). If the decrease in Bdf3 occupancy in differentiating parasites is partly responsible for promoting the transition to a procyclic-specific transcriptome, then an artificial I-BET151 induced decrease in Bdf3 occupancy might also produce the observed changes in transcript levels for procyclic-associated genes.

An unanswered question is whether the transient increase in Bdf3 occupancy observed at pol II initiation sites during differentiation induced by cis-aconitate and temperature shift is also observed during the I-BET151/Bdf3 kd induced differentiation-like phenotype. The earliest timepoint that was used to check for whether I-BET151 decreased binding to Bdf3 targets was 6 hours, and thus a transient increase in occupancy followed by I-BET151 treatment (by some unknown mechanism) might have been missed. Additionally, Bdf3 target occupancy was not tested genome-wide in that previous study (26). It is also possible that the I-BET151 induced differentiation phenotype occurs via an entirely different mechanism than the one induced by cis-aconitate and temperature shift, and that this mechanism does not require a transient increase in Bdf3 occupancy. One model is that the I-BET151 disruption of monoallelic expression is responsible for driving the transcriptome changes that are observed following drug treatment, consistent with studies by Batram et al (64). The idea that differing triggers can result in differentiation phenotypes is also supported by other work (65, 66).

We previously observed an increase in *EP1* transcript levels following bromodomain inhibition (26). This may indicate that Bdf3 occupancy at the *EP1* locus is not required to increase transcript levels of *EP1* or that the increase in *EP1* transcript levels occurs via a different mechanism following bromodomain inhibition, as mentioned above. Alternatively, residual levels of bromodomain protein following inhibition may have been sufficient to drive the observed increase in *EP1.* While ITC experiments confirmed binding of I-BET151 to trypanosome bromodomains (26), it is also possible that I-BET151 has off-target effects that result in the observed increase in EP1 transcript levels.

A number of genes known to be associated with differentiation were within 5kb of a dynamically occupied Bdf3 site (Table S6 lists all genes within 5kb of 'high confidence' Bdf3 sites and separates out those identified by Queiroz et al. as having altered transcript levels during differentiation (67); information from additional studies is included in the third tab of Table S6 (8, 65, 68)). These genes include, but are not limited to, invariant surface glycoproteins (*ISG75*, Tb927.5.350, Tb927.5.360, Tb927.5.370, Tb927.5.380, Tb927.5.390, Tb927.5.400) , procyclin genes on chromosome 6: *GPEET2* gene (Tb927.6.510), *EP3-2* (Tb927.6.520), Procyclin Associated Gene 3 (Tb927.6.530) (Fig. S6), Expression Site Associated Genes (*ESAG1*, Tb927.4.1200, Tb927.10.105, *ESAG2*, Tb927.1.2040, *ESAG3*, Tb927.5.4600, Tb927.9.15940, *ESAG4*, Tb927.5.285b, *ESAG5*, Tb927.7.6860, Tb927.9.15890, *ESAG9*, Tb927.5.4620, Tb927.5.120, Tb927.7.170, *ESAG* pseudogenes, Tb927.1.2060, Tb927.1.2070, Tb927.2.660, Tb927.2.910, Tb927.3.5790, Tb927.9.680, Tb927.9.16010, Tb927.9.16010, Tb927.9.16010,

Tb927.10.100), Variant Surface Glycoproteins (*VSG*s), *COX* genes (Tb927.4.4620, Tb927.9.3170, Tb927.10.8320, Tb927.11.13140), flagellar genes (Tb927.5.4480, Tb927.8.5440, Tb927.8.5460, Tb927.8.5470, and adenylate cyclases (Tb927.11.1480, Tb927.5.320, Tb927.5.330, Tb927.6.270, Tb927.7.7470) (Table S6).

In conclusion, we have adapted the CUT&RUN technique for use in *T. brucei* parasites, and used it to track *de novo* Bdf3 peak formation and changes in occupancy at Bdf3 binding sites during the transition from the bloodstream to the procyclic form. The mechanistic details for how changes in bromodomain protein occupancy might promote differentiation is an exciting area for future study.

Materials and Methods

Parasite Culture and Strain Generation

Bloodstream parasites were cultured in HMI-9 at 37ºC with 5% CO_2. Differentiation was induced by resuspending parasites in Differentiation Media (DTM) at 4 million cells/ml (69), adding 6mM cis-aconitate, and incubating parasites at 27ºC. Post-differentiation parasites were maintained between 1 and 10 million cells/ml. The *BDF3*-HA/*BDF3*KO strain was generated from EATRO 1125 AnTat1.1 90:13 (70). EATRO1125 Antat 1.1 lines were kept at densities below 600,000 cells/ml. The pMOTAG5H *BDF3*-HA construct (26) was linearized and introduced into parasites using an AMAXA nucleofector kit. Correct integration was verified using PCR with primers: 1) HA rev (tatgggtacgcgtaatcaggcaca) and 2) upstream Bdf3 5'UTR for (tgttgcaggatattgtgagtga). After this the pyrFEKO *PAC*/GFP *BDF3* KO construct was linearized and transfected into Bdf3-HA tagged parasites as above and verified with primers: 1) GFP for (ctacaacagccacaaggtctat) and 2) downstream Bdf3 3' UTR rev (aaaccgcaaagtgatgaatgg). Control primers shown in figure are 1) Bdf3 3' UTR for (cttgtagacagcggcatggttgg) and 2) downstream Bdf3 3' UTR rev (aaaccgcaaagtgatgaatgg).

CUT&RUN

All spins prior to permeabilization were performed at 10ºC and 2738g. Spins after the permeabilization step occurred at 10ºC and 4602g. 50-75 million parasites were harvested via centrifugation at 10ºC and washed in 1ml NP buffer containing 0.5mM spermidine, 50mM NaCl, 10mM Tris-HCl pH 7.5, and protease inhibitors. Parasites were spun and permeabilized using 100µl NP buffer supplemented with 0.1% saponin (vol/vol) and 2mM EDTA. 5µg of anti-HA (Sigma H6908) or anti-IgG (Fisher 02-6102) control antibody was added and samples were rotated for 45m at 25ºC. For histone experiments, 1.5µl rabbit anti-H3 (a kind gift from Christian Janzen) was added. Samples were washed and pelleted twice with 1ml NP buffer. A volume corresponding to 1.4% of the sample was removed for flow cytometry analysis. Following the second wash, samples were resuspended in 100µl NP buffer and 0.5µl proteinA-Micrococcal nuclease was added (a kind gift from Steven Henikoff). Samples were rotated for 5m at 25ºC. Samples were washed twice in 1ml NP buffer as above and resuspended in 100µl NP buffer. $CaCl_2$ was added to a final concentration of 2mM and nuclease digestion occurred for 5m at 25ºC. 100µl of 2X STOP buffer (20mM EDTA, 20mM EGTA) with yeast spike-in DNA was immediately added. Samples were incubated for 10m at 37ºC to release insoluble nuclear chromatin. Samples were pelleted and the supernatant containing the DNA was saved. SDS was added to a final concentration of 0.1%, proteinase K was added at 165µg/ml, and RNaseA was added to 6.5µg/ml. Samples were incubated at 70ºC for 10m and purified using phenol chloroform or Ampure XP beads at 1.8X according to the manufacturer's instructions.

Flow Cytometry

All flow cytometry was performed on a Novocyte 2000R from Acea Biosciences (now Agilent). Parasites were resuspended in 100µl HMI-9 and stained for 10 minutes on ice with mouse anti-rabbit IgG PE (Santa Cruz sc-3753). Cells were washed twice in HMI-9 prior to analysis.

Quantitative PCR Analysis

To quantify transcript levels of *PAD1*, parasites were grown to a density of 190,000 cells/ml for low density samples or >= 1 million cells/ml for high density samples. RNA was extracted from low or high density parasite populations using RNA Stat-60 (Tel-Test) following the manufacturer's protocol and quantified on a NanoDrop2000c. 2.5µg of RNA was used to generate cDNA using the SuperScript IV VILO Master Mix (Fisher Scientific 11756050) according to the manufacturer's protocol. cDNA was amplified using 2X Sybr green master mix (Life Technologies 4309155) and primers and quantified on an Eppendorf Realplex2 instrument. Primers used were *PAD1*: gaccaaaggaaccttcttcct and cactggctcccctaagct, *URA3*: cggcagcagttctcgagt and tggcgtgtaccttgaggc. For differentiation experiments, primers used for *EP1* were tctgctcgctattcttctgttc and cctttgcctcccttagtaagac, and *Tb927.10.9400 SF1* were ggtatggttcatcaggagttgg and cgttagcactggtatccttcag.

Generation of Sequencing Libraries

Sequencing libraries were generated using the NEBNext Ultra II DNA Library Prep Kit for Illumina (E7645) according to the manufacturer's instructions with the following modification: Samples were incubated with USER enzyme immediately prior to PCR, rather than at an earlier step. NEBNext Multiplex Oligos for Illumina were used to prepare multiplex samples (e.g. E7710).

CUT&RUN Sequencing Analysis

Sequencing was performed at the UCLA Technology Center for Genomics and Bioinformatics using an Illumina HISEQ 3000 with 50bp single end reads. CUT&RUN fastq files were trimmed using TrimGalore (http://www.bioinformatics.babraham.ac.uk/projects/trim_galore/) and aligned to the Tb927v5.1 genome using bowtie (71) and requiring unique alignments using the following command: `bowtie --best --strata -t -v 2 -a -m 1`. Spike-in reads were aligned to the yeast sacCer3/R64 genome. MACS (42) was used in broad peak mode to identify peaks of Bdf3 localization using an IgG control with the following arguments
`-g 23650671, --keep-dup all, --nomodel, and –broad.`
The DiffBind (48) package was used to identify regions with a change in Bdf3 occupancy. The GreyList ChIP package eliminated problematic regions from the analysis (0.66% of the genome). Four different normalization methods were then used to obtain normalized read counts for areas of interest determined by MACS: Spike-in library size, Spike-in RLE, Background RLE (not using spike-in reads), and RPKM. Both RLE methods adjust for regional 'background' read frequency by counting reads in non-overlapping 15kb genomic bins. This approach adjusts for broad patterns in background read enrichment while avoiding spurious adjustment to locally enriched regions. When MACS regions are used as input for the DiffBind program, it defines regions of interest as xbp upstream and downstream of the summit for the peak determined by MACS, where x is a variable determined by the user. We tested a number of x values and found similar results between 100 and 500 basepairs. The analysis presented here is performed with 200bp on either side of the summit; thus regions of interest are each 400bp long. DiffBind was used to identify regions with a statistically significant change in read counts (occupancy) at each time point using bloodstream values as a control with a cutoff of $p_{adj} < 0.05$ using the Benjamini-Hochberg adjustment to control the FDR per normalization

technique. Once these regions were identified using each normalization method, a Venn diagram was used to identify 'high confidence' regions with changes in occupancy that were identified using all four normalization methods. To generate control regions, we randomly shuffled 'high confidence' regions for each chromosome using bedtools (56). Shuffled regions were maintained on the same chromosome and peaks of Bdf3 localization identified by MACS were excluded. Control regions were run through the same DiffBind pipeline to obtain normalized read counts, which were plotted for each control region.

Overlap between ChIP and CUT&RUN-identified peaks was performed by first merging peaks that were within 5kb in each dataset and then identifying unique vs overlapping peaks using bedtools (56).

De novo peak formation was ascertained using a custom script. The program took as input the set of MACS called Bdf3 consensus peaks at each time point and identified sites that were not called as Bdf3 peaks in bloodstream forms but subsequently were called as peaks at 2 consecutive timepoints following differentiation. The script for 'disappearing' peaks took the same input and identified sites that were called as peaks in bloodstream parasites and not called as peaks at two consecutive time points following differentiation.

Data Availability

Fastq files are deposited at the SRA under project number PRJNA795567.

Acknowledgements

We would like to thank Christian Janzen for his kind gift of the anti-H3 antibody. We also thank Monica Mugnier for helpful comments on the manuscript. This work was funded by a NSF CAREER Grant (2041395) to Danae Schulz.

Figure 1. **Optimization of CUT&RUN for *T. brucei***. A) Flow cytometry analysis for parasites permeabilized with 0.1% Saponin and incubated with rabbit anti-H3 primary antibody and then stained with anti-Rabbit PE. Left, forward and side scatter with live gate shown. Right, live gated parasites stained with anti-H3 (blue) or processed in parallel but with no anti-H3 added and stained with secondary antibody alone (red). B) Gel electrophoresis of DNA isolated from CUT&RUN processed parasites using an anti-H3 antibody. $CaCl_2$ was added for the indicated times and temperatures to activate micrococcal nuclease cutting. C) Gel electrophoresis of DNA isolated from CUT&RUN processed parasites using an anti-H3 antibody (left panels and first lane on right panel) or an IgG antibody (right panel, second lane). Left panel shows samples processed with or without $CaCl_2$ to activate micrococcal nuclease cutting. Black line indicates a gel cropped to show the indicated samples.

Figure 2. Generation of pleomorphic *BDF3*-HA/*BDF3*KO parasites. A) Schematic of both endogenous *Bdf3* alleles following modification with a knockout construct (top) and an HA-tagging construct (bottom). Primers used to assay for correct integration are indicated. B) Electrophoresis of genomic DNA amplified with primers to verify correct integration of the *BDF3* knockout construct (Primers 1 and 2, top panel), primers to verify correct integration of the *BDF3*-HA tagging construct (Primers 3 and 4, middle panel), and control primers that should amplify all genomic DNA samples (Primers 5 and 2, bottom panel). Transformant clones are indicated as numbers. HA strain no KO indicates strain modified with the HA construct but not the KO construct. Control strains are two non-pleomorphic strains used as negative controls for the modifications. White line indicates cropping of gel photo. C) Quantitative PCR experiment to assay the transcript levels of *PAD1* and a control gene (*URA3)* in Bdf3-HA tagged pleomorphic

parasites at low density (less than 400,000 cells/ml) or high density (800-1000000 cells/ml). Values for the *rRNA* transcript were used for normalization. Each dot represents the average of 3 technical replicates. 3 biological replicates were used for the data shown here. The long horizontal bar represents the average value for the 3 biological replicates. Error bars represent the standard error. The data was scaled such that the average for the low density samples was set to 1 to make it easier to compare different gene targets. * indicates p-value < 0.05 for a Student's unpaired T test with equal variance comparing *Pad1* expression at low density vs high density.

Figure 3. **CUT&RUN identified Bdf3 binding sites are similar to those previously found by ChIP-seq.** A) IGV display for a region of chromosome 7 showing sequencing tracks for five biological replicates processed for CUT&RUN using an anti-HA antibody in a pleomorphic Bdf3-HA tagged parasite line. A control sample processed with anti-IgG is also shown. Blue boxes below sequencing tracks indicate peaks of Bdf3 localization called by MACS in CUT&RUN samples and in published Bdf3-HA ChIP experiments in monomorphic strains (26). The last row displays a gene track. B) Plot showing the number of Bdf3 ChIP-identified peaks that do or do not overlap with Bdf3 CUT&RUN-identified peaks and vice versa. Materials and methods describe details of peak merging for this analysis.

Figure 4. Most Bdf3 binding sites are retained throughout differentiation. IGV display for a region of chromosome 9 showing sequencing tracks for overlaid biological replicates processed for CUT&RUN using an anti-HA antibody in a pleomorphic Bdf3-HA tagged parasite line. A control sample processed with anti-IgG is also shown. Blue boxes below each sequencing track indicate peaks of Bdf3 localization identified by MACS.

Figure 5. *De novo* Bdf3 peaks appear at the *EP1* locus after differentiation to the procyclic form is induced. Top panel. IGV display for a region of chromosome 10 showing sequencing tracks for overlaid biological replicates processed for CUT&RUN using an anti-HA antibody in a pleomorphic Bdf3-HA tagged parasite line. A control sample processed with anti-IgG is also shown. Blue boxes below each sequencing track indicate peaks of Bdf3 localization identified by MACS. The black square outlines the location of a new peak forming at the *EP1* locus. Bottom panel. Same as top panel except zoomed in on the *EP1* locus for chromosome 10. Three peaks are called at 76 hours that did not appear previously. Green box in front of *EP1* indicates the promoter. Arrowhead indicates direction of transcription.

Figure 6. Changes in occupancy for Bdf3 are found at many loci throughout the genome following induction of differentiation from the bloodstream to the procyclic form. Venn diagram showing the number of sites identified by DiffBind as having a statistically significant change in occupancy when compared to bloodstream samples using four different normalization methods.

Figure 7. Bdf3 occupancy peaks at 3 hours after induction of differentiation. A) Plots of $\log_2$ normalized tag counts over time for each 'high confidence' Bdf3 binding site with a statistically significant change in occupancy after induction of differentiation to the procyclic form using background RLE normalization and arrayed by chromosome. B) Plots of $\log_2$ normalized tag counts over time for control regions that are not identified as Bdf3 binding sites using background RLE normalization.

Supplemental Figure 1. **The *EP* locus shows a different pattern of occupancy compared to loci within 5kb of differentiation-associated genes.** A) Plot showing $\log_2$ RLE normalized

counts for Bdf3 occupancy at the 3 *de novo* peaks detected at the *EP* locus. Plot in the right panel includes individual values for each replicate. B) Plot showing $\log_2$ RLE normalized counts for Bdf3 occupancy at sites within 5kb of genes known to be associated with differentiation.

Supplemental Figure 2. *EP1* expression increases at 3h post differentiation in pleomorphic parasites. A) Quantitative PCR experiment to assay the transcript levels of *EP1* and a control gene (*SF1)* in Antat 1.1 pleomorphic parasites following differentiation at the indicated timepoints. Values for the *rRNA* transcript were used for normalization. Each dot represents the average of 3 technical replicates. 3 biological replicates were used for the data shown here. The horizontal bar represents the average value for the 3 biological replicates. Error bars represent the standard error. The data was scaled such that the average for the 0h samples was set to 1 to make it easier to compare different gene targets. * indicates p-value < 0.05 for a Student's unpaired T test with equal variance comparing *EP1* expression at indicated time point vs 0h timepoint.
B) Same as in A except using the Bdf3-HA tagged pleomorphic line. Each dot represents the average of 3 technical replicates. 3 biological replicates were used for the data shown here. The horizontal bar represents the average value for the 3 biological replicates. Error bars represent the standard error. The data was scaled such that the average for the 0h samples was set to 1 to make it easier to compare different gene targets. * indicates p-value < 0.05 for a Student's unpaired T test with equal variance comparing *EP1* expression at indicated time point vs 0h timepoint.

Supplemental Figure 3. IGV display for a region of chromosome 5 showing sequencing tracks for overlaid biological replicates processed for CUT&RUN using an anti-HA antibody in a pleomorphic Bdf3-HA tagged parasite line. A control sample processed with anti-IgG is also shown. Blue boxes below each sequencing track indicate peaks of Bdf3 localization identified by MACS. A) *De novo* peak at a locus near a putative *ISG* (Tb927.5.309b). B) *De novo* peak near a cluster of *ISG75* genes (Tb927.5.350-380).

Supplemental Figure 4. Plots of $\log_2$ normalized tag counts over time for each 'high confidence' Bdf3 binding site with a statistically significant change in occupancy after induction of differentiation to the procyclic form using indicated normalization.

Supplemental Figure 5. Plots of $\log_2$ normalized tag counts over time for control regions that are not identified as Bdf3 binding sites using indicated normalization.

Supplemental Figure 6. IGV display for a region of chromosome 6 at the *GPEET* (Tb927.6.530) locus showing sequencing tracks for overlaid biological replicates processed for CUT&RUN using an anti-HA antibody in a pleomorphic Bdf3-HA tagged parasite line. A control sample processed with anti-IgG is also shown. Blue boxes below each sequencing track indicate peaks of Bdf3 localization identified by MACS.

Supplemental Table 1. Diffbind normalized tag counts over time for Bdf3 peaks identified by MACS. Normalization is performed 4 different ways, and the normalization method is indicated by the tab name in the excel document.

Supplemental Table 2. Analysis of Bdf3 binding sites that appear *de novo* during differentiation.

Supplemental Table 3. Analysis of Bdf3 binding sites that are present in the bloodstream form and not called by MACS following differentiation.

Supplemental Table 4. Normalized tag counts over time for Bdf3 sites with dynamic occupancy identified by DiffBind using four normalization methods.

Supplemental Table 5. Normalized tag counts over time for control regions that are not Bdf3 binding sites using four normalization methods. Normalization is performed 4 different ways, and the normalization method is indicated by the tab name in the excel document.

Supplemental Table 6. Genes within 5kb of a 'high confidence' dynamically occupied Bdf3 site.

Supplemental Table 7. Divergent strand switch regions used for this study.

Supplemental Table 8. Consensus sites for Bdf3 binding regions for all time points with gene overlaps and TSS overlaps indicated.

Supplemental Table 9. Consensus sites for Bdf3 binding regions for all time points.

References

1. Capewell P, Cren-Travaillé C, Marchesi F, Johnston P, Clucas C, Benson RA, Gorman T-A, Calvo-Alvarez E, Crouzols A, Jouvion G, Jamonneau V, Weir W, Stevenson ML, O'Neill K, Cooper A, Swar NK, Bucheton B, Ngoyi DM, Garside P, Rotureau B, MacLeod A. 2016. The skin is a significant but overlooked anatomical reservoir for vector-borne African trypanosomes. eLife 5:e17716.

2. Trindade S, Rijo-Ferreira F, Carvalho T, Pinto-Neves D, Guegan F, Aresta-Branco F, Bento F, Young SA, Pinto A, Van Den Abbeele J, Ribeiro RM, Dias S, Smith TK, Figueiredo LM. 2016. Trypanosoma brucei Parasites Occupy and Functionally Adapt to the Adipose Tissue in Mice. Cell Host and Microbe 19:837–848.

3. Roditi I, Schumann G, Naguleswaran A. 2016. Environmental sensing by African trypanosomes. Current Opinion in Microbiology 32:26–30.

4. Cross GAM, Kim H-S, Wickstead B. 2014. Capturing the variant surface glycoprotein repertoire (the VSGnome) of Trypanosoma brucei Lister 427. Molecular & Biochemical Parasitology 195:59–73.

5. Hovel-Miner G, Mugnier MR, Papavasiliou FN, Pinger J, Schulz D. 2015. A Host-Pathogen Interaction Reduced to First Principles: Antigenic Variation in T. brucei. Results Probl Cell Differ 57:23–46.

6. Rico E, Rojas F, Mony BM, Szöőr B, MacGregor P, Matthews KR. 2013. Bloodstream form pre-adaptation to the tsetse fly in Trypanosoma brucei. Front Cell Infect Microbiol 3:78.

7. Vassella E, Acosta-Serrano A, Studer E, Lee SH, Englund PT, Roditi I. 2001. Multiple procyclin isoforms are expressed differentially during the development of insect forms of Trypanosoma brucei. J Mol Biol 312:597–607.

8. Briggs EM, Rojas F, McCulloch R, Matthews KR, Otto TD. 2021. Single-cell transcriptomic analysis of bloodstream Trypanosoma brucei reconstructs cell cycle progression and developmental quorum sensing. Nat Commun 12:5268.

9. Jensen BC, Sivam D, Kifer CT, Myler PJ, Parsons M. 2009. Widespread variation in transcript abundance within and across developmental stages of Trypanosoma brucei. BMC Genomics 10:482.

10. Kabani S, Fenn K, Ross A, Ivens A, Smith TK, Ghazal P, Matthews K. 2009. Genome-wide expression profiling of in vivo-derived bloodstream parasite stages and dynamic analysis of mRNA alterations during synchronous differentiation in Trypanosoma brucei. BMC Genomics 10:427.

11. Nilsson D, Gunasekera K, Mani J, Osteras M, Farinelli L, Baerlocher L, Roditi I, Ochsenreiter T. 2010. Spliced Leader Trapping Reveals Widespread Alternative Splicing Patterns in the Highly Dynamic Transcriptome of Trypanosoma brucei. PLoS Pathog 6:e1001037.

12. Cayla M, McDonald L, MacGregor P, Matthews K. 2020. An atypical DYRK kinase connects quorum-sensing with posttranscriptional gene regulation in Trypanosoma brucei. Elife 9:e51620.

13. Dean S, Marchetti R, Kirk K, Matthews KR. 2009. A surface transporter family conveys the trypanosome differentiation signal. Nature 459:213–217.

14. McDonald L, Cayla M, Ivens A, Mony BM, MacGregor P, Silvester E, McWilliam K, Matthews KR. 2018. Non-linear hierarchy of the quorum sensing signalling pathway in bloodstream form African trypanosomes. PLoS Pathog 14:e1007145.

15. Mony BM, MacGregor P, Ivens A, Rojas F, Cowton A, Young J, Horn D, Matthews K. 2014. Genome-wide dissection of the quorum sensing signalling pathway in Trypanosoma brucei. Nature 505:681–685.

16. Reuner B, Vassella E, Yutzy B, Boshart M. 1997. Cell density triggers slender to stumpy differentiation of Trypanosoma brucei bloodstream forms in culture. Mol Biochem Parasitol 90:269–280.

17. Rojas F, Silvester E, Young J, Milne R, Tettey M, Houston DR, Walkinshaw MD, Pérez-Pi I, Auer M, Denton H, Smith TK, Thompson J, Matthews KR. 2019. Oligopeptide Signaling through TbGPR89 Drives Trypanosome Quorum Sensing. Cell 176:306-317.e16.

18. Vassella E, Reuner B, Yutzy B, Boshart M. 1997. Differentiation of African trypanosomes is controlled by a density sensing mechanism which signals cell cycle arrest via the cAMP pathway. Journal of Cell Science 110 (Pt 21):2661–2671.

19. Jha BA, Gazestani VH, Yip CW, Salavati R. 2015. The DRBD13 RNA binding protein is involved in the insect-stage differentiation process of Trypanosoma brucei. FEBS Lett 589:1966–1974.

20. Kolev NG, Ramey-Butler K, Cross GAM, Ullu E, Tschudi C. 2012. Developmental Progression to Infectivity in Trypanosoma brucei Triggered by an RNA-Binding Protein. Science 338:1352–1353.

21. Mugo E, Clayton C. 2017. Expression of the RNA-binding protein RBP10 promotes the bloodstream-form differentiation state in Trypanosoma brucei. PLoS Pathog 13:e1006560.

22. Mugo E, Egler F, Clayton C. 2017. Conversion of procyclic-form Trypanosoma brucei to the bloodstream form by transient expression of RBP10. Molecular & Biochemical Parasitology 216:49–51.

23. Paterou A, Walrad P, Craddy P, Fenn K, Matthews K. 2006. Identification and stage-specific association with the translational apparatus of TbZFP3, a CCCH protein that promotes trypanosome life-cycle development. J Biol Chem 281:39002–39013.

24. Walrad P, Paterou A, Acosta-Serrano A, Matthews KR. 2009. Differential Trypanosome Surface Coat Regulation by a CCCH Protein That Co-Associates with procyclin mRNA cis-Elements. PLoS Pathog 5:e1000317.

25. Walrad PB, Capewell P, Fenn K, Matthews KR. 2012. The post-transcriptional trans-acting regulator, TbZFP3, co-ordinates transmission-stage enriched mRNAs in Trypanosoma brucei. Nucleic Acids Res 40:2869–2883.

760 26. Schulz D, Mugnier MR, Paulsen E-M, Kim H-S, Chung CW, Tough DF, Rioja I, Prinjha RK,
761 Papavasiliou FN, Debler EW. 2015. Bromodomain Proteins Contribute to Maintenance of
762 Bloodstream Form Stage Identity in the African Trypanosome. PLoS Biol 13:e1002316-38.

763 27. Dhalluin C, Carlson JE, Zeng L, He C, Aggarwal AK, Zhou M-M, Zhou M-M. 1999. Structure
764 and ligand of a histone acetyltransferase bromodomain. Nature 399:491–496.

765 28. Jenuwein T, Allis CD. 2001. Translating the Histone Code. Science 293:1074–1080.

766 29. Siegel TN, Hekstra DR, Kemp LE, Figueiredo LM, Lowell JE, Fenyo D, Wang X, Dewell S,
767 Cross GAM. 2009. Four histone variants mark the boundaries of polycistronic transcription
768 units in Trypanosoma brucei. Genes & Development 23:1063–1076.

769 30. Yang X, Wu X, Zhang J, Zhang X, Xu C, Liao S, Tu X. 2017. Recognition of hyperacetylated
770 N-terminus of H2AZ by TbBDF2 from Trypanosoma brucei. Biochem J 474:3817–3830.

771 31. Zaware N, Zhou M-M. 2019. Bromodomain biology and drug discovery. Nat Struct Mol Biol
772 26:870–879.

773 32. Zeng L, Yap KL, Ivanov AV, Wang X, Mujtaba S, Plotnikova O, Rauscher FJ, Zhou M-M.
774 2008. Structural insights into human KAP1 PHD finger-bromodomain and its role in gene
775 silencing. Nat Struct Mol Biol 15:626–633.

776 33. Winter GE, Mayer A, Buckley DL, Erb MA, Roderick JE, Vittori S, Reyes JM, di Iulio J,
777 Souza A, Ott CJ, Roberts JM, Zeid R, Scott TG, Paulk J, Lachance K, Olson CM, Dastjerdi
778 S, Bauer S, Lin CY, Gray NS, Kelliher MA, Churchman LS, Bradner JE. 2017. BET
779 Bromodomain Proteins Function as Master Transcription Elongation Factors Independent
780 of CDK9 Recruitment. Mol Cell 67:5-18.e19.

781 34. Di Micco R, Fontanals-Cirera B, Low V, Ntziachristos P, Yuen SK, Lovell CD, Dolgalev I,
782 Yonekubo Y, Zhang G, Rusinova E, Gerona-Navarro G, Cañamero M, Ohlmeyer M,
783 Aifantis I, Zhou M-M, Tsirigos A, Hernando E. 2014. Control of Embryonic Stem Cell
784 Identity by BRD4-Dependent Transcriptional Elongation of Super-Enhancer-Associated
785 Pluripotency Genes. Cell Reports 9:234–247.

786 35. Horne GA, Stewart HJS, Dickson J, Knapp S, Ramsahoye B, Chevassut T. 2015. Nanog
787 requires BRD4 to maintain murine embryonic stem cell pluripotency and is suppressed by
788 bromodomain inhibitor JQ1 together with Lefty1. Stem Cells Dev 24:879–891.

789 36. Najafova Z, Tirado-Magallanes R, Subramaniam M, Hossan T, Schmidt G, Nagarajan S,
790 Baumgart SJ, Mishra VK, Bedi U, Hesse E, Knapp S, Hawse JR, Johnsen SA. 2017.
791 BRD4 localization to lineage-specific enhancers is associated with a distinct transcription
792 factor repertoire. Nucleic Acids Res 45:127–141.

793 37. Lee J-E, Park Y-K, Park S, Jang Y, Waring N, Dey A, Ozato K, Lai B, Peng W, Ge K. 2017.
794 Brd4 binds to active enhancers to control cell identity gene induction in adipogenesis and
795 myogenesis. Nat Commun 8:2217.

796 38. Staneva DP, Carloni R, Auchynnikava T, Tong P, Rappsilber J, Jeyaprakash AA, Matthews
797 KR, Allshire RC. 2021. A systematic analysis of *Trypanosoma brucei* chromatin factors
798 identifies novel protein interaction networks associated with sites of transcription initiation
799 and termination. Genome Res 31:2138–2154.

800 39. Skene PJ, Henikoff S. 2017. An efficient targeted nuclease strategy for high-resolution
801 mapping of DNA binding sites. eLife 6:e21856.

40. Teytelman L, Thurtle DM, Rine J, van Oudenaarden A. 2013. Highly expressed loci are vulnerable to misleading ChIP localization of multiple unrelated proteins. Proceedings of the National Academy of Sciences 110:18602–18607.

41. Kornberg RD. 1974. Chromatin Structure: A Repeating Unit of Histones and DNA. Science 184:868–871.

42. Zhang Y, Liu T, Meyer CA, Eeckhoute J, Johnson DS, Bernstein BE, Nusbaum C, Myers RM, Brown M, Li W, Liu XS. 2008. Model-based analysis of ChIP-Seq (MACS). Genome Biol 9:R137.

43. Kaya-Okur HS, Wu SJ, Codomo CA, Pledger ES, Bryson TD, Henikoff JG, Ahmad K, Henikoff S. 2019. CUT&Tag for efficient epigenomic profiling of small samples and single cells. Nat Commun 10:1930.

44. Janssens DH, Meers MP, Wu SJ, Babaeva E, Meshinchi S, Sarthy JF, Ahmad K, Henikoff S. 2021. Automated CUT&Tag profiling of chromatin heterogeneity in mixed-lineage leukemia. Nat Genet 53:1586–1596.

45. Wu SJ, Furlan SN, Mihalas AB, Kaya-Okur HS, Feroze AH, Emerson SN, Zheng Y, Carson K, Cimino PJ, Keene CD, Sarthy JF, Gottardo R, Ahmad K, Henikoff S, Patel AP. 2021. Single-cell CUT&Tag analysis of chromatin modifications in differentiation and tumor progression. Nat Biotechnol 39:819–824.

46. Waldman BS, Schwarz D, Wadsworth MH, Saeij JP, Shalek AK, Lourido S. 2020. Identification of a Master Regulator of Differentiation in Toxoplasma. Cell 180:359-372.e16.

47. Koenig-Martin E, Yamage M, Roditi I. 1992. A procyclin-associated gene in Trypanosoma brucei encodes a polypeptide related to ESAG 6 and 7 proteins. Mol Biochem Parasitol 55:135–145.

48. Stark, Rory B Gord. 2011. DiffBind: Differential binding analysis of ChIP-Seq peak data. http://bioconductor.org/packages/release/bioc/vignettes/DiffBind/inst/doc/DiffBind.pdf.

49. Brown SD, Huang J, Van der Ploeg LH. 1992. The promoter for the procyclic acidic repetitive protein (PARP) genes of Trypanosoma brucei shares features with RNA polymerase I promoters. Mol Cell Biol 12:2644–2652.

50. Gunzl A, Bruderer T, Laufer G, Schimanski B, Tu LC, Chung HM, Lee PT, Lee MGS. 2003. RNA Polymerase I Transcribes Procyclin Genes and Variant Surface Glycoprotein Gene Expression Sites in Trypanosoma brucei. Eukaryotic Cell 2:542–551.

51. Rudenko G, Lee MG, Van der Ploeg LH. 1992. The PARP and VSG genes of Trypanosoma brucei do not resemble RNA polymerase II transcription units in sensitivity to Sarkosyl in nuclear run-on assays. Nucleic Acids Res 20:303–306.

52. Shea C, Lee MG, Van der Ploeg LH. 1987. VSG gene 118 is transcribed from a cotransposed pol I-like promoter. Cell 50:603–612.

53. Clayton CE, Fueri JP, Itzhaki JE, Bellofatto V, Sherman DR, Wisdom GS, Vijayasarathy S, Mowatt MR. 1990. Transcription of the procyclic acidic repetitive protein genes of Trypanosoma brucei. Mol Cell Biol 10:3036–3047.

54. Biebinger S, Rettenmaier S, Flaspohler J, Hartmann C, Peña-Diaz J, Wirtz LE, Hotz HR, Barry JD, Clayton C. 1996. The PARP promoter of Trypanosoma brucei is developmentally regulated in a chromosomal context. Nucleic Acids Res 24:1202–1211.

55. Ross-Innes CS, Stark R, Teschendorff AE, Holmes KA, Ali HR, Dunning MJ, Brown GD, Gojis O, Ellis IO, Green AR, Ali S, Chin S-F, Palmieri C, Caldas C, Carroll JS. 2012. Differential oestrogen receptor binding is associated with clinical outcome in breast cancer. Nature 481:389–393.

56. Quinlan AR, Hall IM. 2010. BEDTools: a flexible suite of utilities for comparing genomic features. Bioinformatics 26:841–842.

57. Domingo-Sananes MR, Szöőr B, Ferguson MAJ, Urbaniak MD, Matthews KR. 2015. Molecular control of irreversible bistability during trypanosome developmental commitment. The Journal of Cell Biology 211:455–468.

58. McKnight JN, Boerma JW, Breeden LL, Tsukiyama T. 2015. Global Promoter Targeting of a Conserved Lysine Deacetylase for Transcriptional Shutoff during Quiescence Entry. Molecular Cell 59:732–743.

59. Swygert SG, Kim S, Wu X, Fu T, Hsieh T-H, Rando OJ, Eisenman RN, Shendure J, McKnight JN, Tsukiyama T. 2019. Condensin-Dependent Chromatin Compaction Represses Transcription Globally during Quiescence. Molecular Cell 73:533-546.e4.

60. Young CP, Hillyer C, Hokamp K, Fitzpatrick DJ, Konstantinov NK, Welty JS, Ness SA, Werner-Washburne M, Fleming AB, Osley MA. 2017. Distinct histone methylation and transcription profiles are established during the development of cellular quiescence in yeast. BMC Genomics 18:107.

61. Cucinotta CE, Dell RH, Braceros KC, Tsukiyama T. 2021. RSC primes the quiescent genome for hypertranscription upon cell-cycle re-entry. eLife 10:e67033.

865 62. Jones NG, Geoghegan V, Moore G, Carnielli JBT, Newling K, Calderón F, Gabarró R,
866 Martín J, Prinjha R, Rioja I, Wilkinson AJ, Mottram JC. 2021. Bromodomain factor 5 is an
867 essential transcriptional regulator of the *Leishmania* genome. preprint. Cell Biology.

868 63. Evans C, Hardin J, Stoebel DM. 2018. Selecting between-sample RNA-Seq normalization
869 methods from the perspective of their assumptions. Brief Bioinform 19:776–792.

870 64. Batram C, Jones NG, Janzen CJ, Markert SM, Engstler M. 2014. Expression site
871 attenuation mechanistically links antigenic variation and development in Trypanosoma
872 brucei. eLife 3:e02324.

873 65. Qiu Y, Milanes JE, Jones JA, Noorai RE, Shankar V, Morris JC. 2018. Glucose Signaling Is
874 Important for Nutrient Adaptation during Differentiation of Pleomorphic African
875 Trypanosomes. mSphere 3:e00366-18.

876 66. Quintana JF, Zoltner M, Field MC. 2021. Evolving Differentiation in African Trypanosomes.
877 Trends Parasitol 37:296–303.

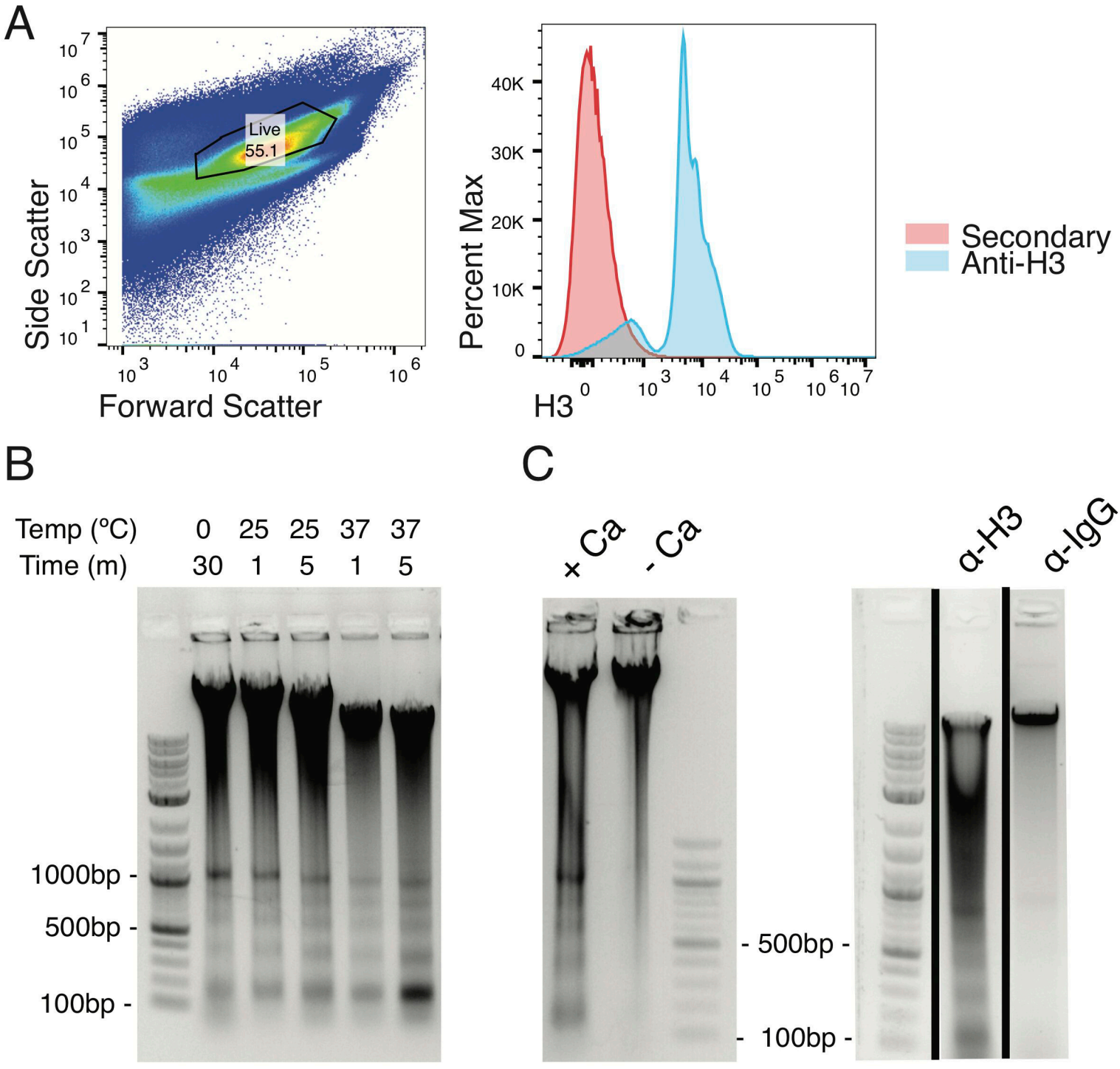
878 67. Queiroz R, Benz C, Fellenberg K, Hoheisel JD, Clayton C. 2009. Transcriptome analysis of
879 differentiating trypanosomes reveals the existence of multiple post-transcriptional regulons.
880 BMC Genomics 10:495.

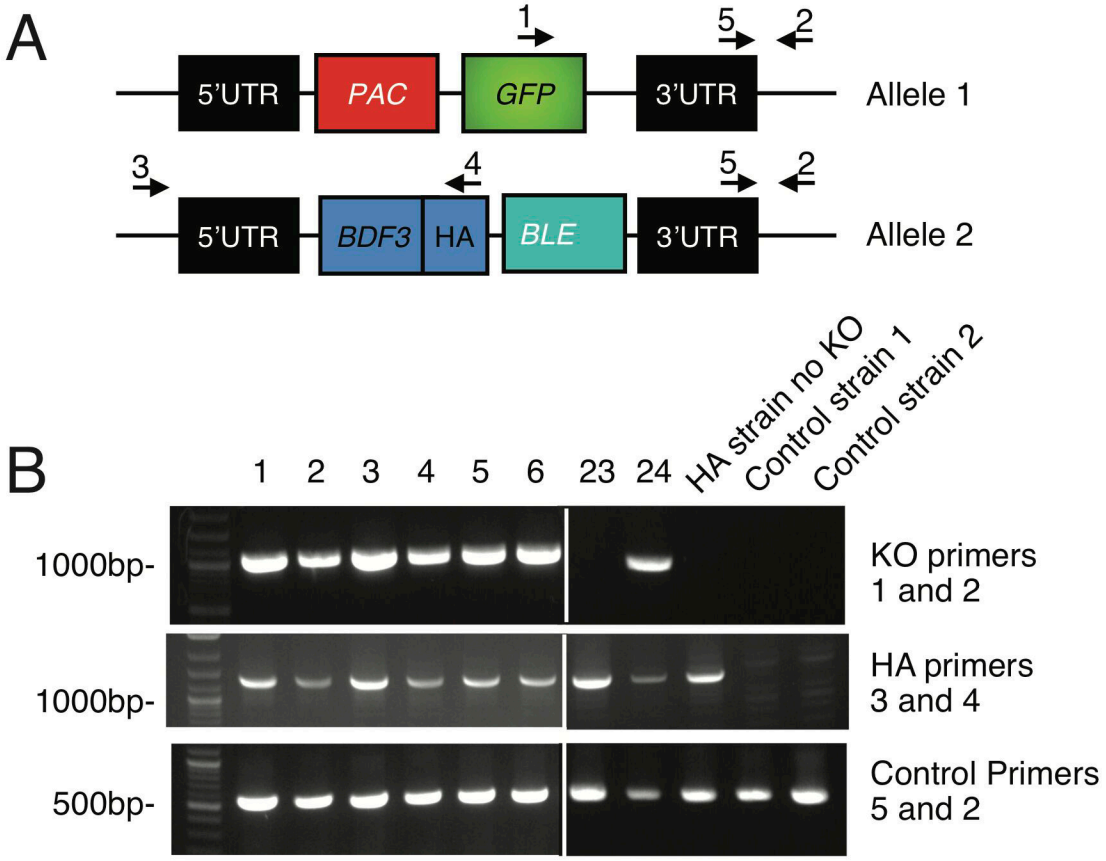
881 68. Siegel TN, Hekstra DR, Wang X, Dewell S, Cross GAM. 2010. Genome-wide analysis of
882 mRNA abundance in two life-cycle stages of Trypanosoma brucei and identification of
883 splicing and polyadenylation sites. Nucleic Acids Research 38:4946–4957.

884 69. Ziegelbauer K, Quinten M, Schwarz H, Pearson TW, Overath P. 1990. Synchronous
885 differentiation of Trypanosoma brucei from bloodstream to procyclic forms in vitro. Eur J
886 Biochem 192:373–378.

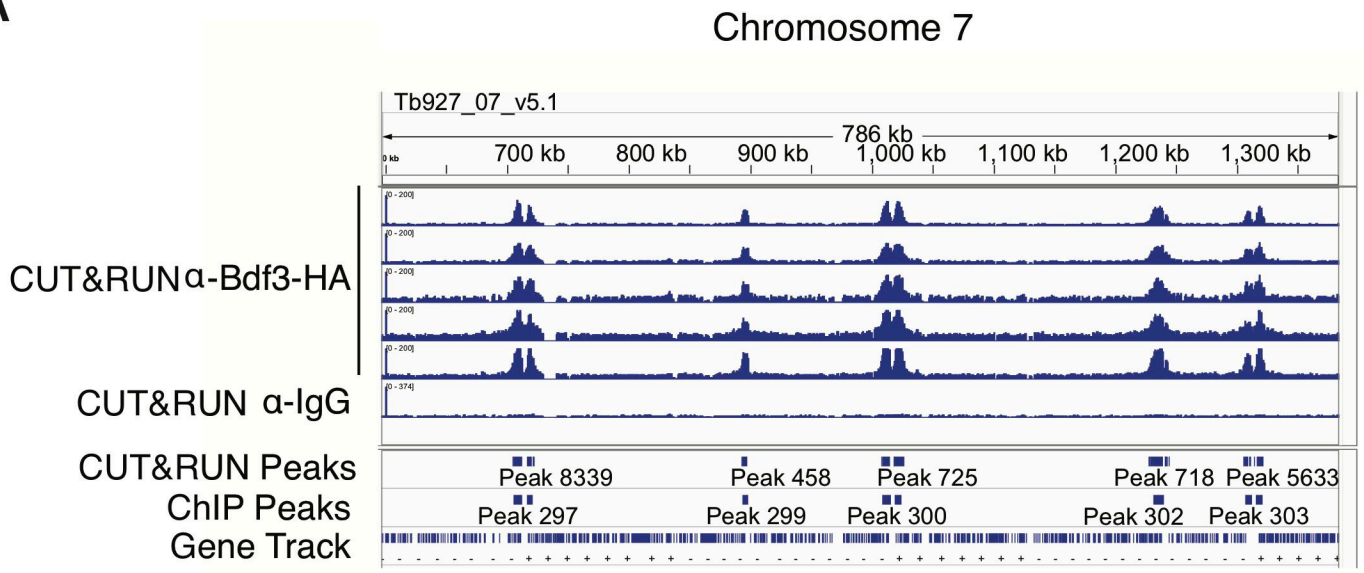
70. Engstler M, Boshart M. 2004. Cold shock and regulation of surface protein trafficking convey sensitization to inducers of stage differentiation in Trypanosoma brucei. Genes Dev 18:2798–2811.

71. Langmead B, Trapnell C, Pop M, Salzberg SL. 2009. Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. Genome Biol 10:R25.

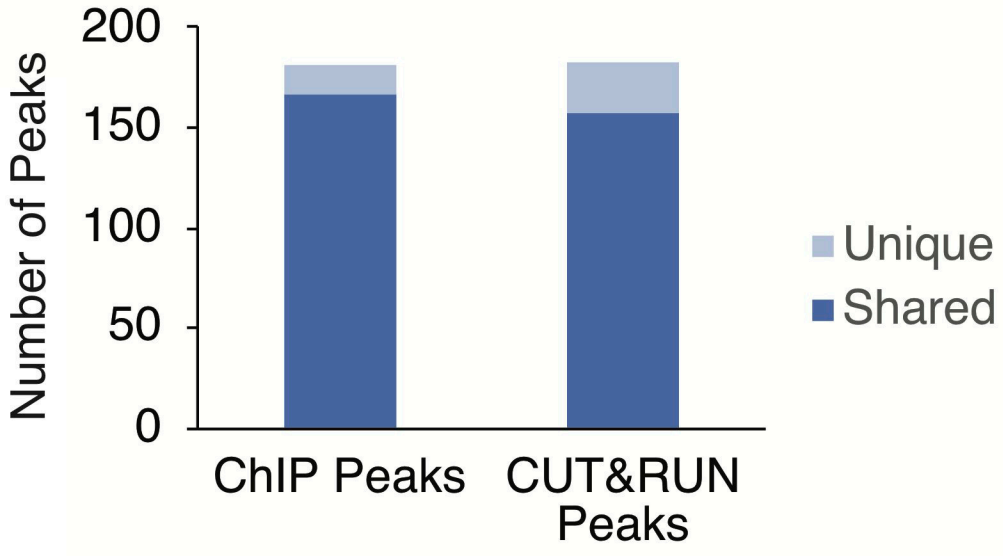

A
Side Scatter
Forward Scatter
Live
55.1
Percent Max
H3
Secondary
Anti-H3
B
Temp (°C) 0 25 25 37 37
Time (m) 30 1 5 1 5
1000bp -
500bp -
100bp -
C
+ Ca
- Ca
α-H3
α-IgG
- 500bp -
- 100bp -

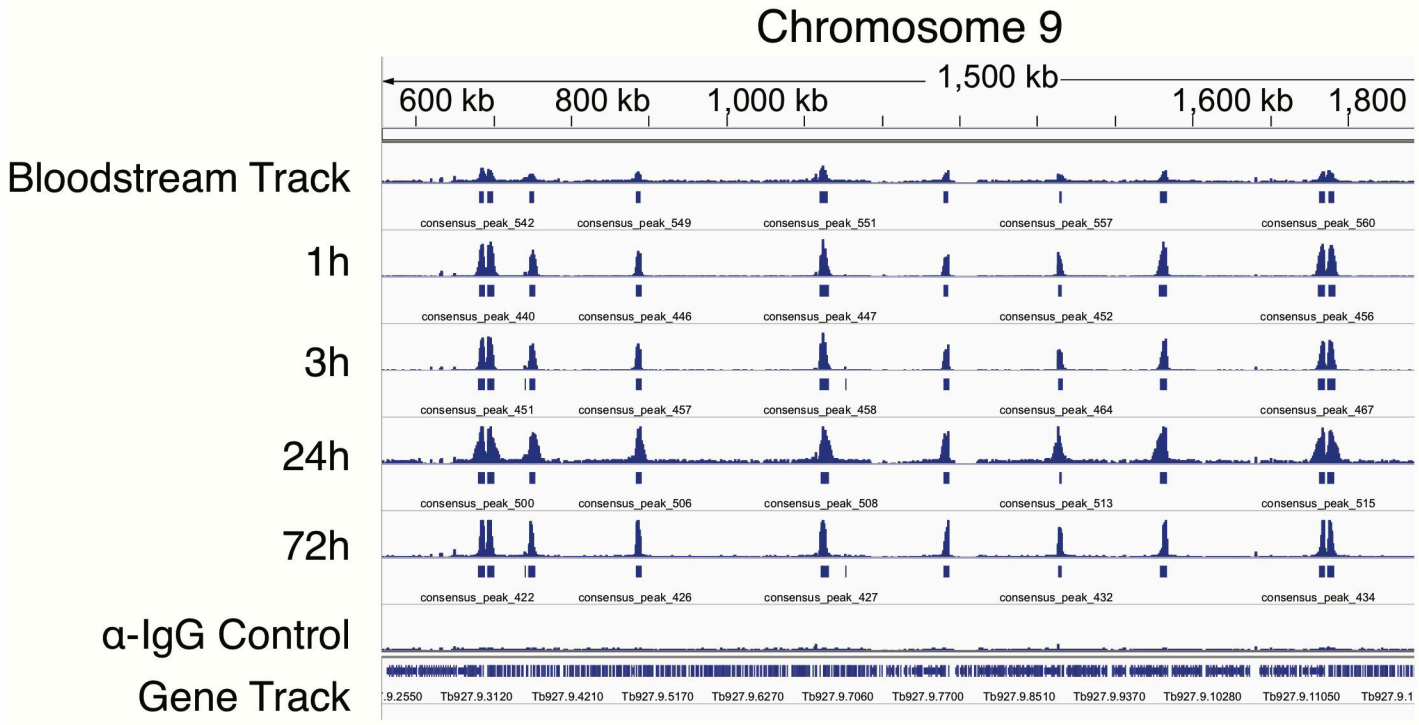

A
1
5
2
5'UTR
PAC
GFP
3'UTR
Allele 1
3
4
5
2
5'UTR
BDF3
HA
BLE
3'UTR
Allele 2
B
1 2 3 4 5 6 23 24
HA strain no KO
Control strain 1
Control strain 2
1000bp-
KO primers
1 and 2
1000bp-
HA primers
3 and 4
500bp-
Control Primers
5 and 2
C
relative expression
0
2
4
6
8
10
Low Density
High Density
*
Pad1
Ura3

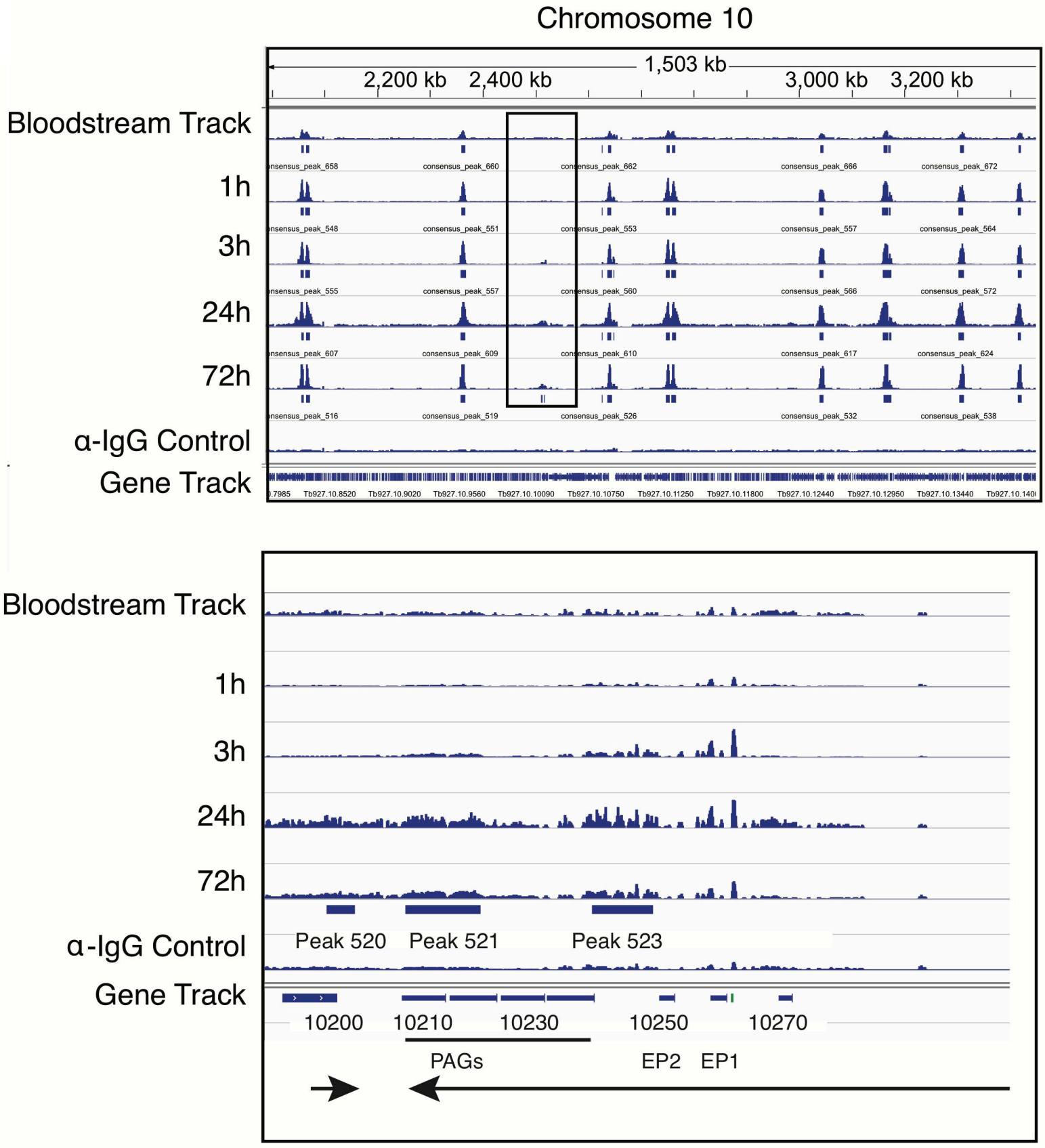
A

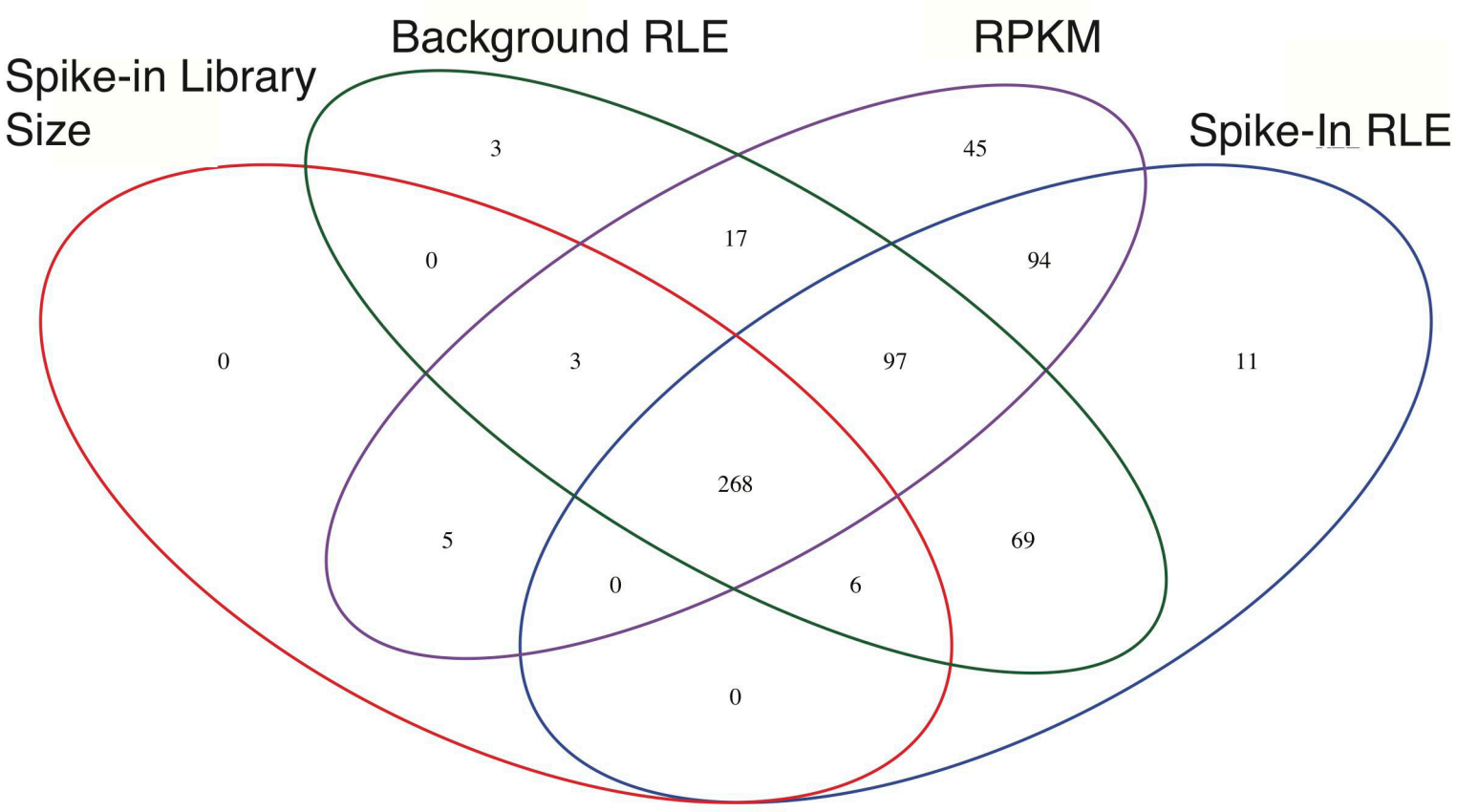

Chromosome 7
Tb927_07_v5.1
786 kb
700 kb
800 kb
900 kb
1,000 kb
1,100 kb
1,200 kb
1,300 kb
CUT&RUN α-Bdf3-HA
CUT&RUN α-IgG
CUT&RUN Peaks
ChIP Peaks
Gene Track
Peak 8339
Peak 458
Peak 725
Peak 718
Peak 5633
Peak 297
Peak 299
Peak 300
Peak 302
Peak 303

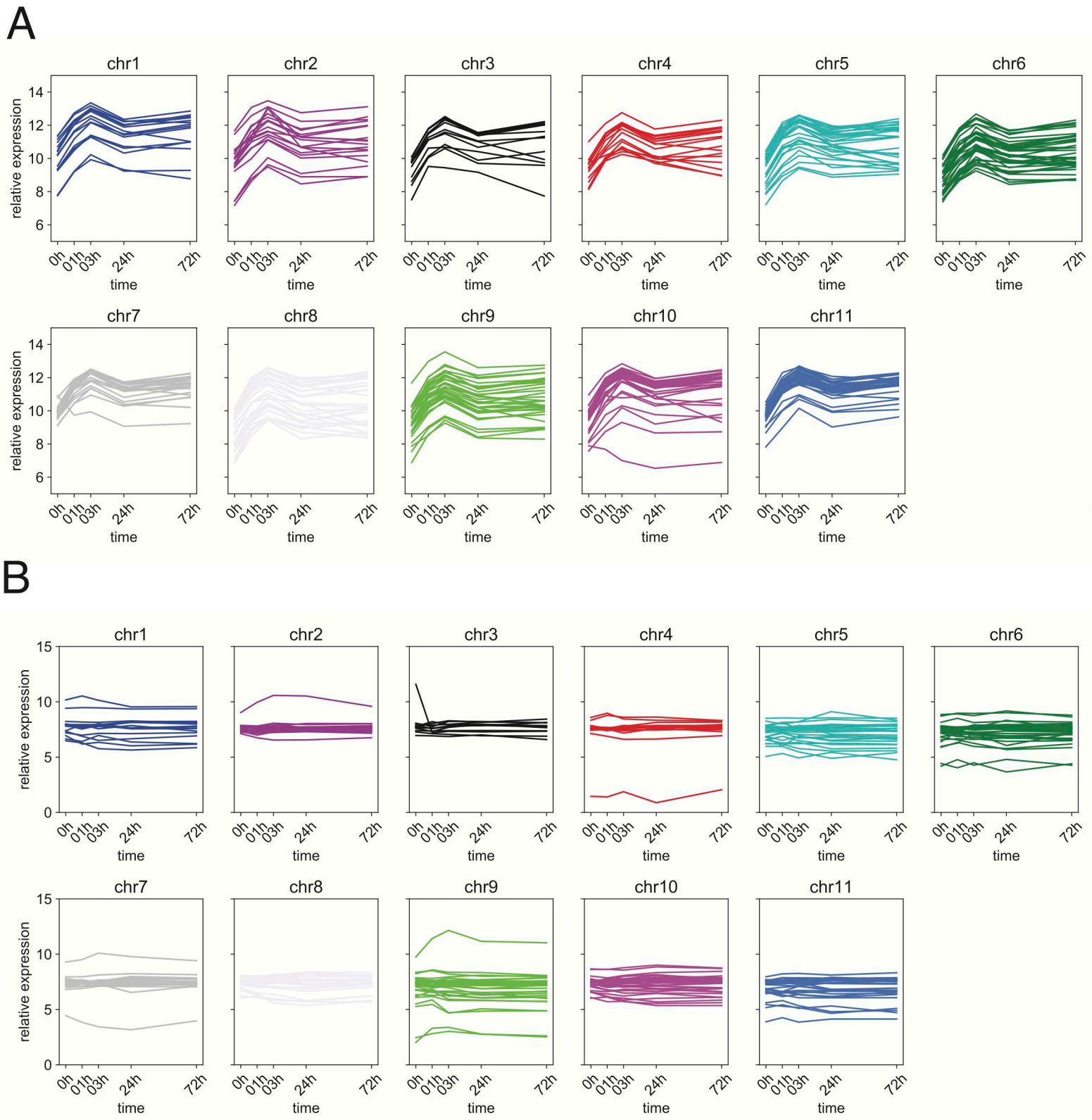
B


Number of Peaks
200
150
100
50
0
ChIP Peaks
CUT&RUN Peaks
Unique
Shared


Chromosome 9
600 kb
800 kb
1,000 kb
1,500 kb
1,600 kb
1,800
Bloodstream Track
consensus_peak_542
consensus_peak_549
consensus_peak_551
consensus_peak_557
consensus_peak_560
1h
consensus_peak_440
consensus_peak_446
consensus_peak_447
consensus_peak_452
consensus_peak_456
3h
consensus_peak_451
consensus_peak_457
consensus_peak_458
consensus_peak_464
consensus_peak_467
24h
consensus_peak_500
consensus_peak_506
consensus_peak_508
consensus_peak_513
consensus_peak_515
72h
consensus_peak_422
consensus_peak_426
consensus_peak_427
consensus_peak_432
consensus_peak_434
α-IgG Control
Gene Track
.9.2550
Tb927.9.3120
Tb927.9.4210
Tb927.9.5170
Tb927.9.6270
Tb927.9.7060
Tb927.9.7700
Tb927.9.8510
Tb927.9.9370
Tb927.9.10280
Tb927.9.11050
Tb927.9.1


Chromosome 10
1,503 kb
2,200 kb
2,400 kb
3,000 kb
3,200 kb
Bloodstream Track
1h
3h
24h
72h
α-IgG Control
Gene Track
consensus_peak_658
consensus_peak_660
consensus_peak_662
consensus_peak_666
consensus_peak_672
consensus_peak_548
consensus_peak_551
consensus_peak_553
consensus_peak_557
consensus_peak_564
consensus_peak_555
consensus_peak_557
consensus_peak_560
consensus_peak_566
consensus_peak_572
consensus_peak_607
consensus_peak_609
consensus_peak_610
consensus_peak_617
consensus_peak_624
consensus_peak_516
consensus_peak_519
consensus_peak_526
consensus_peak_532
consensus_peak_538
Tb927.10.8520
Tb927.10.9020
Tb927.10.9560
Tb927.10.10090
Tb927.10.10750
Tb927.10.11250
Tb927.10.11800
Tb927.10.12440
Tb927.10.12950
Tb927.10.13440
Bloodstream Track
1h
3h
24h
72h
α-IgG Control
Gene Track
Peak 520
Peak 521
Peak 523
10200
10210
10230
10250
10270
PAGs
EP2
EP1


Spike-in Library Size
Background RLE
RPKM
Spike-In RLE
3
45
0
17
94
0
3
97
11
268
5
69
0
6
0


A
chr1
chr2
chr3
chr4
chr5
chr6
chr7
chr8
chr9
chr10
chr11
relative expression
time
0h
01h
03h
24h
72h
B
chr1
chr2
chr3
chr4
chr5
chr6
chr7
chr8
chr9
chr10
chr11
relative expression
time